Circular dichroism analysis for multidomain proteins: studies of the irreversible unfolding of *Hepatitis C virus* helicase

Agnieszka Gozdek, Anna Stankiewicz-Drogoń, Jarosław Poznański✉
and Anna M. Boguszewska-Chachulska✉

Institute of Biochemistry and Biophysics PAS, Warszawa, Poland

The non-structural protein 3 (NS3) of *Hepatitis C virus* (HCV) is a bifunctional enzyme with RNA-dependent NTPase/RNA helicase and serine protease activities, and thus represents a promising target for anti-HCV therapy. These functions are performed by two distinct moieties; the N-terminal protease domain and the C-terminal helicase domain that further folds into three structural subdomains. To obtain lower molecular mass proteins suitable for nuclear magnetic resonance studies of helicase-inhibitor complexes, helicase domains 1, 2, and 1+2 devoid of a hydrophobic β-loop were overexpressed and purified. Circular dichroism studies were carried out to confirm the secondary structure content and to determine thermodynamic parameters describing the stability of the proteins. Both thermal and GuHCl-induced unfolding experiments confirmed the multidomain organization of the helicase. The unfolding transition observed for domain 1+2 was in agreement with the model of two well-resolved successive steps corresponding to the independent unfolding of domains 1 and 2, respectively. In the case of the full-length helicase, the presence of domain 3 remarkably changed the transition profile, leading to fast and irreversible transformation of partially unfolded protein.

INTRODUCTION

Hepatitis C virus (HCV) is a monopartite single-stranded (+)RNA virus, the type member of the *Hepacivirus* genus in the *Flaviviridae* family. It represents a serious worldwide epidemiological threat, infecting up to 3% of the population (180 million; WHO, 2007). Chronic hepatitis resulting from HCV infection develops in most cases into cirrhosis and highly increases the risk of cancer. There are neither fully efficient methods of treatment nor vaccines; even the new dual therapy with peginterferon and ribavirin is effective only in up to 63% of the cases (Hadziyannis *et al.*, 2004). One of the most promising targets for anti-HCV therapy seems to be the viral non-structural protein 3 (NS3) because of its multi-

ple enzymatic activities: RNA-stimulated nucleoside triphosphatase (NTPase), RNA helicase and serine protease (Kwong *et al.*, 2000; Bretner, 2005; Gordon & Keller, 2005). The helicase activity is indispensable for viral replication *in vivo*, presumably unwinding double-stranded replication intermediates or regions, allowing RNA amplification; it is also important for the synthesis of the viral polyprotein (Rosenberg, 2001).

The HCV NS3 helicase is composed of three separate structural domains (Fig. 1): domain 1 containing NTPase motifs; domain 2 with several conserved motifs, among them an arginine-rich sequence involved in ATP binding (Yao *et al.*, 1997; Kwong *et al.*, 2000); and the highly helical domain 3 with conserved residues (Trp-501 and Glu-493) that

✉Corresponding authors: Anna M. Boguszewska-Chachulska; Jarosław Poznański, Institute of Biochemistry and Biophysics PAS, A. Pawińskiego 5a, 02-106 Warszawa, Poland; fax: (48) 22 658 4636; e-mail: annach@ibb.waw.pl (A.M.B.-C.); jarek@ibb.waw.pl (J.P.).
Abbreviations: CD, circular dichroism; GuHCl, guanidine hydrochloride; HCV, Hepatitis C virus; ITC, isothermal titration colorimetry; MS, mass spectrometry; NMR, nuclear magnetic resonance; NS3, non-structural protein 3; NTPase, nucleoside triphosphatase.
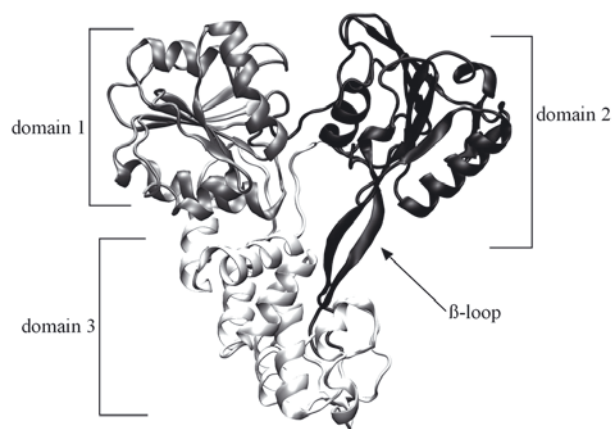
Figure 1. Ribbon diagram of the three-domain HCV NS3 helicase (PDB entry 1A1V).

The hydrophobic β-loop removed from the domain 1+2 construct is indicated by an arrow.

are confined only to helicases from HCV and related viruses (Lam *et al.*, 2003). No thermodynamic parameters describing the helicase stability are available since up to now there has been only one study concerning the stability of the HCV helicase-derived proteins, namely domain 2 in the presence of double-stranded DNA (Liu *et al.*, 2003), and one on the thermodynamic parameters of the HCV helicase interaction with DNA (Levin & Patel, 2002).

Using relaxation filtered nuclear magnetic resonance (NMR) experiments, binding of inhibitors to the full-length HCV helicase can be detected (Sarver *et al.*, 2002), but in spite of recent developments of NMR equipment, it is still necessary to study individual domains separately to precisely locate sites of interaction. Since the HCV helicase aggregates at high concentrations and in some experimental conditions (Levin & Patel, 2002; and personal experience), we looked for a more stable and non-aggregating fragment of the protein that could mimic the biochemical properties of the entire enzyme. Within the frame of our project aimed at selecting efficient helicase inhibitors we decided to clone the full-length HCV helicase and its separate domains, as well as a construct covering its two most important domains, 1 and 2, for NMR, fluorometric and isothermal titration calorimetry (ITC) studies on the interaction of the protein with previously identified anti-helicase compounds (Boguszewska-Chachulska *et al.*, 2006).

To check whether proteins expressed in various cell systems are properly folded and thus can be used in further structural experiments, circular dichroism (CD) technique is used, although its application for multidomain proteins is still difficult (Kelly & Price, 1997). The NMR experiments planned also include the determination of thermodynamic parameters of inhibitor binding to the helicase domains

and for this purpose the stability of the separate domains must be evaluated. The CD approach is a method of choice also for this type of analysis.

Here we report the preparation of appropriate helicase constructs and results of CD analyses performed to evaluate the structural stability of the helicase and its subcloned domains, prerequisites for further structural studies of this multidomain protein.

MATERIALS AND METHODS

Expression and purification of proteins in the baculovirus system. The expression and purification of the full-length HCV helicase and its domain 2 were preformed as previously described (Boguszewska-Chachulska *et al.*, 2004).

Cloning and expression of domain 1 of the HCV helicase. Domain 1 was cloned in the bifunctional YpET-30a vector, a recombinant of pET-30a (Novagen) and pFL38, a kind gift of Dr. M. Zagulski (Institute of Biochemistry and Biophysics PAS), using homologous recombination in the *Saccharomyces cerevisiae* BY strain. This method relies on plasmid gap-repair by PCR-generated cassettes (Zagulski *et al.*, 2003). Oligonucleotide primers were obtained from the Laboratory of DNA Sequencing and Oligonucleotide Synthesis (Institute of Biochemistry and Biophysics PAS). HCV-1 cDNA template (Acc. No. AAA45676) (Kim *et al.*, 1995) was used to amplify domain 1 with the sense primer (GTTTAACTTTAAGAAGGAGATATACATA-TG**GTGGACTTTATCCCTGTGGAG**) and the antisense primer (CGGATCTCAGTGGTGGTGGTGGTG-GTG**AGGGGTGGCGGTGGCGAGC**) that covered the His$_6$-tag coding sequence from the vector (nucleotides corresponding to the sequence of domain 1 are in bold). Total DNA was isolated from yeast, recombinant plasmids were identified by PCR, and amplified in *Escherichia coli* DH5α. The correct sequence of the construct was confirmed by restriction digestion and sequencing. Protein overexpression was carried out in the *E. coli* BL21(DE3)LysE strain in Luria–Bretani medium supplemented with kanamycin (50 μg/ml). Cells grown at 37°C until the cultures reached an OD$_{600}$ of approx. 1.0 were induced with 1 mM isopropyl-β-ᴅ-thiogalactopyranoside (IPTG) and harvested after 4 h of incubation at 37°C. The recombinant protein was purified according to the protocol established for proteins purified from insect cell cultures (Boguszewska-Chachulska *et al.*, 2004), with some modifications. The pellet from 1 liter of culture was resuspended in 10 ml of buffer A (20 mM Tris/HCl, pH 7.0, 500 mM NaCl, and 5 mM imidazole). The cells were frozen in liquid nitrogen, brought to 37°C and vortexed; freezing and thawing

were repeated five times. The suspension was centrifuged at 20 000 × g at 4°C for 15 min. Subsequently, glycerol was added to the clear supernatant to 10% and Chaps to 0.05%, the extract was mixed with 1 ml of Talon Metal Affinity Resin (Clontech) washed previously with 10 ml of buffer B (buffer A with 10% glycerol and 0.05% Chaps), and left overnight on a rocking platform at 4°C. The mixture was washed four times with 10 ml of buffer B, and then three times with 10 ml of buffer A. Five elutions with 1 ml of buffer E (20 mM Tris/HCl, pH 7.0, 200 mM NaCl and 200 mM imidazole) were performed. The concentration of HCV helicase domain 1 was determined at 280 nm using the molar absorption coefficient of 7680 M^{-1} cm^{-1}. The absorption coefficient was calculated on the basis of protein composition by the ProtParam program from the Expasy website (http://us.expasy.org).

Cloning and expression of domain 1+2 of the HCV helicase. The initial domain 1+2 construct in YpET-30a was obtained as described for domain 1. The PCR/recombination primers used were: sense GTTTAACTTTAAGAAGGAGATATA-CATATG**GTGGACTTTATCCCTGTGGAG** and antisense CGGATCTCAGTGGTGGTGGTGGTGGTG-*GGAGGGGCGCTCACCCGGTGC* (nucleotides corresponding to the sequence of domain 1 are in bold, those corresponding to domain 2 are in italics). To enhance protein solubility two outer β-sheet fragments of domain 2 (amino acids 256–286) were removed, as described by Gesell *et al.* (2001). The construct containing this deletion was obtained by PCR amplification of the YpET-30a plasmid including the entire sequence of domain 1+2 using appropriate primers (sense AAA<u>ggatcc</u>**AAG**CCCCAAGAT-GCTGTCTCC, antisense CTT<u>ggatcc</u>**GTCGGA**CG-TATTGCAGTCTA; nucleotides corresponding to the domain 2 sequence are in bold; nucleotides corresponding to the *Bam*HI site are in lower case and underlined) introducing a sequence coding for a loop (SDGSK) to replace the two removed β-sheet fragments (CVTQTVDFSLDPTFTIETITL). The PCR product, digested with *Bam*HI and re-circularised with T4 DNA ligase (Invitrogen), was used to transform *E. coli* DH5α. Recombinant clones containing the newly introduced *Bam*HI site were isolated and sequenced. Selected plasmids were used to transform *E. coli* BL21(DE3)LysE. The modified protein was expressed and purified according to the protocol developed for domain 1. Its concentration was determined using the molar absorption coefficient of 14 440 M^{-1} cm^{-1}. All the experiments with domain 1+2 presented here were carried out using the domain 1+2 construct devoid of the hydrophobic loop.

Protein analysis. Extracts and eluates were analyzed by 12 or 17% SDS/PAGE and proteins visualized by Coomassie staining or electrotransferred onto an Immobilon-P membrane (Millipore). Proteins were detected with rabbit polyclonal anti-HCV helicase antibodies (obtained at the State Veterinary Institute, Puławy, Poland), followed by a goat anti-rabbit IgG (whole molecule)-alkaline phosphatase conjugate (Sigma) as secondary antibodies. The immunoreactive proteins were revealed colorimetrically using chromogenic substrates (NBT/BCIP Solution, Roche).

CD measurements. Initial stock solutions of the HCV helicase and its domains were prepared in 50 mM Na_iPO_4 buffer, pH 7.6, and 0.3 M NaCl. Far UV-CD spectra were measured in a cell with a 2 mm path-length at 200–270 nm with an AVIV 202 spectropolarimeter, and the thermal denaturation of domain 1+2 was carried out using a JASCO J-810 spectropolarimeter. The experimental parameters were the following: temperature 25°C; bandwidth 1 nm; wavelength step 1 nm; averaging time 1 s. CD spectra analyses were performed for the full-length helicase, its separate domains 1 and 2, and the construct corresponding to domains 1+2.

Measurements performed at various temperatures were carried out in 50 mM Na_iPO_4 buffer, pH 7.6, and 0.3 M NaCl in a 300 μl cell with a 2 mm optical path-length. The temperature varied in the range of 30 to 80°C with a gradient of 2°C/min. The unfolding process was monitored in 2°C steps at 220 nm with an averaging time of 3 s. Refolding was induced by temperature decrease to 30°C with a gradient of 2°C/min.

Titration analysis was performed by stepwise replacement of 10 μl of the protein solution in the sample cell by 6 M denaturant (guanidine hydrochloride, GuHCl) in 50 mM Na_iPO_4 buffer, pH 7.6, with 0.3 M NaCl. Since the molar ellipticity of the HCV helicase was independent of the protein concentration, this type of titration experiment was used to reduce the amount of protein required. Five independent titration experiments starting with 0, 1, 2, 3 and 4 M GuHCl were performed with the number of replacements sufficient to obtain a significant overlap of concentration ranges in subsequent experiments. Finally, spectra were measured in 6 M GuHCl as a reference for the unfolded protein.

Analysis of CD spectra. In order to judge whether the protein fragments were properly folded, the partition of secondary structure elements was estimated with the help of the CDNN program (Bohm *et al.*, 1992) based on a neural network trained on a set of 23 proteins. Due to intensive scattering in the short wavelength region, the structural analyses were performed on the basis of data obtained for the restricted spectral range of 200–270 nm.

Modelling of partially irreversible thermal denaturation. Numerical analysis of the thermal denaturation of the protein was based on the three-state Lumry–Eyring model of unfolding (Lumry & Eyring, 1954; Frost & Pearson, 1961) describing the folding-

unfolding equilibrium in the presence of an irreversibly unfolded form, according to the general scheme:

$$N \xleftrightarrow{K} U \xrightarrow{k_f} F \tag{1}$$

where N and U are the native and unfolded states experiencing the equilibrium characterized by the middle point temperature, T_U, and the free energy change upon the transition, ΔG^U, according to the relation:

$$\frac{[U]}{[N]} = K = \frac{x_U}{x_N} = e^{-\Delta G^U (1/RT - 1/RT_U)} \tag{2}$$

Bearing in mind that the transition to a final irreversibly unfolded state, F, undergoes efficiently only in the relatively narrow temperature range in which the reversibly unfolded state, U, is significantly populated (cf. Eqn. 4), the temperature dependence of the rate of the U→F transition, k_f, may be neglected:

$$k_f = k_o \tag{3}$$

As it was additionally verified, the estimated changes of the rate k_f varied within a narrow range of approx. 10% for experimental temperatures ranging from 30 to 80°C. The above assumption did not decrease the quality of the mathematical model, significantly improving the stability of the optimization procedure used. Finally, the concentration balance involving three types of species is described by the following system of Eqns:

$$x_N + x_U + x_F = 1$$
$$x_U = x_N * e^{-\Delta G^U (1/RT - 1/RT_U)}$$
$$\frac{dx_F}{dt} = x_U * k_f \tag{4}$$

where x_i is partition of the i-th form of the protein. Bearing in mind that the unfolding experiment is performed by discrete step-wise changes of temperature (2°C jump followed by 1 min equilibration), and the reversible unfolding is characterized by the equilibrium constant K, the above Eqns could be formally separated into time-dependent and time-independent Eqns:

$$\frac{dx_F}{dt} = k_f * (1 - x_F) \frac{K}{1+K}$$
$$x_N = (1 - x_F) \frac{1}{1+K}$$
$$x_U = (1 - x_F) \frac{K}{1+K} \tag{5}$$

where K and k_f are defined in Eqns 2 and 3, respectively. Assuming that the initial partition of the irreversibly unfolded form at the beginning of the equilibration period performed at temperature T equals $x_F(T,0)$, the time-dependent part of Eqn. 5 could be formally integrated, giving the general formula:

$$x_F(T,t) = 1 - [1 - x_F(T,0)] * e^{-t \cdot k_f \cdot \frac{K}{1+K}} \tag{6}$$

Eqn. (6) is to be summed over all the discrete temperature changes of the system, while K is defined by Eqn. (2) and k_f is defined by Eqn. (3). The system response function, θ(T), e.g. 222 nm ellipticity measured after 1 min equilibration at temperature T, was then analyzed as the function resulting from the sequence of preprogrammed changes of the sample temperature, according to the formula:

$$\theta(T) = (a_N T + \theta_N) \times x_N(T) + (a_U T + \theta_U) \times x_U(T) + C \tag{7}$$

where θ_N and θ_U are the ellipticity of the native and unfolded states, respectively, and a_N, a_U and C serve as the first order correction for the θ(T) relationship. The thermodynamic parameters: ΔG^U, T_U and k_o, together with the CD-related parameters: θ_N, θ_U, a_N, a_U and C were then optimized to obtain the best agreement between the simulated and experimental data. To improve the numerical stability of the algorithm used, the three types of thermal unfolding experiments were performed separately, but subsequently globally analyzed upon single optimization.

The standard experiment of thermal unfolding was enhanced by additional experiments in which the temperature was decreased after reaching the assumed level, T_R, and, when needed, the sample was additionally incubated for a given time, τ_R, at the temperature T_R. For every sample the set of T_R values and possible usage of τ_R delays were rationalized individually. Generally, at least three different T_R values were chosen in the 2–3°C range of the apparent middle point of the N ↔ U transition recorded in the standard experiment. The possible delay, τ_R, was set to additional 5 min following 1 min equilibration at the temperature T_R.

Modelling GuHCl-induced unfolding. Analysis of the GuHCl titration experiments was based on the model of a sequence of independent two-state transitions. For an individual transition the partition of the folded protein, x_N, was assumed in the form:

$$x_N(T) = \left(1 + e^{-\frac{\Delta G^U + m[GuHCl]}{RT}}\right)^{-1} \tag{8}$$

where ΔG^U is a measure of the protein stability in denaturant-free solution, T is the temperature, [GuHCl] is the concentration of the unfolding agent, and m is the efficiency of the unfolding agent. Multistate transitions were modelled as the superposition of independent two-state processes.

When the partition of the folded structure was estimated from CD spectroscopic measurements at 222 nm as a function of either temperature or denaturing agent concentration, the model parameters were optimized using the Marquardt–Levenberg non-linear least-squares algorithm (Marquardt, 1963) implemented in Gnuplot 4.1 (Williams & Kelley, 2007).

RESULTS AND DISCUSSION

Overexpression and purification of the HCV helicase and its domains

The HCV helicase and domain 2 were purified for structural studies with a yield of 15 mg helicase and 20 mg domain 2 per liter of culture, as described previously (Boguszewska-Chachulska *et al.*, 2004). To facilitate cloning of domain 1 that remained very difficult in spite of various approaches applied, we used a bifunctional vector, YpET-30a, constructed by recombination, where the autonomous replication/centromeric sequence (ARS-CEN) region and a selection marker gene complementing uracil auxotrophy (*URA3*) from pFL38 were inserted into pET-30a. This vector made immediate cloning of domain 1 and domain 1+2 possible by homologous recombination in yeast.

Domain 1 and domain 1+2 were expressed as C-terminally His$_6$-tagged fusion proteins in *E. coli*. Non-denaturing cell lysis was followed by native purification on Talon Resin that resulted in 95% purity as estimated by SDS/PAGE and Coomassie staining (Fig. 2) with 10 and 9 mg/l of protein, respectively.

The lysates and purified proteins were analyzed by Western blot using polyclonal rabbit antibodies against the NS3 helicase (Fig. 3). The results obtained confirmed the identity of the proteins and
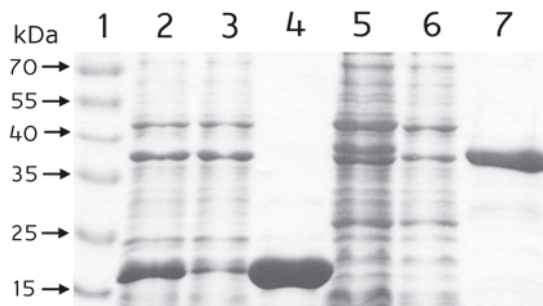

Figure 2. Expression and purification of helicase domains analyzed by SDS/PAGE and detected by Coomassie staining.
Lane 1, prestained molecular markers (SM0671, Fermentas); lane 2, lysate of domain 1; lane 3, flow-through of domain 1; lane 4, domain 1 purified on Talon resin; lane 5, lysate of domain 1+2; lane 6, flow-through of domain 1+2; lane 7, domain 1+2 purified on Talon resin.
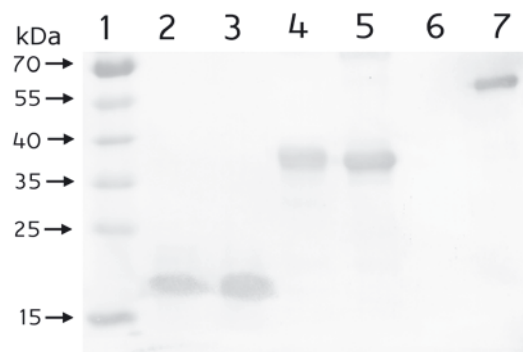

Figure 3. Western blot analysis of expression of domains 1 and 1+2.
Lane 1, prestained molecular markers (SM0671, Fermentas); lane 2, lysate of domain 1; lane 3, purified domain 1; lane 4, lysate of domain 1+2; lane 5, purified domain 1+2; lane 6, negative control (BL21 lysate); lane 7, positive control (full-length HCV helicase).

showed that both proteins carried helicase-specific epitopes. The size of the proteins verified by mass spectrometry (MS) showed that the recombinant proteins had the expected molecular masses of 17.3 and 32.4 kDa, respectively. Thus the use of a recombinant vector and homologous recombination in yeast makes it possible to overcome frequent problems with cloning in protein expression vectors and facilitates further modifications of cloned constructs, such as deletions of fragments of coding sequences. It should be stressed, however, that several clones must be analyzed and sequenced because a significant level of mutations was observed when this cloning method was applied.

Structural analysis of purified proteins

CD experiments were carried out with 3–15 µM helicase-derived proteins. All the proteins studied, the full-length HCV helicase, its separate domains 1 and 2, and the construct encompassing domains 1 and 2, displayed CD spectra adequate for the expected partitions of the secondary structure elements (Fig. 4), clearly indicating that all the proteins analyzed were initially properly folded.

Thermal resistance of proteins

The thermal stability of the HCV helicase appeared relatively low since thermal unfolding of the enzyme was initiated at 35ºC, and above 60ºC the partially unfolded protein precipitated. Primary numerical analyses of the process showed that the main transition occurs at the temperature $T_{app} = 54.8$ºC (Table 1). At higher temperatures the protein aggregates, exhibiting residual ellipticity; this indicates that some protein regions still preserve their
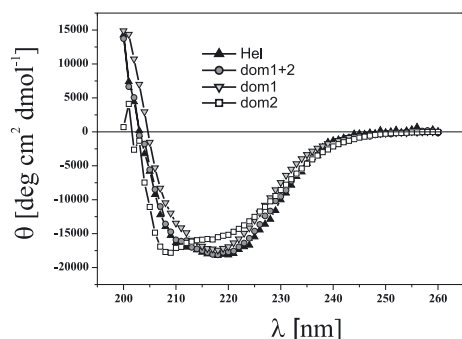
Figure 4. Far-UV CD spectra of the HCV helicase, its separate domains 1 and 2, and domain 1+2.
The spectra agree with the partitions of secondary structure elements deduced from the known 3D structure of the HCV helicase.

helical structure. Since no ellipticity recovery was observed upon cooling of the protein from 80 to 30°C, the helicase thermal denaturation process must be regarded as irreversible. The simplest thermodynamic model of an irreversible unfolding transition proposed by Lumry and Eyring (1954) assumes that the unfolded protein, U, may be irreversibly transferred within the first order reaction U → F to a final state, F. Taking into account that the parameters of the Lumry–Eyring model fitted to the standard unfolding experiment were found strongly correlated, four additional experiments were performed. They encompassed partial thermal unfolding of the protein, followed by cooling. The protein sample was heated to 60°C and then cooled to 30°C. Since this experiment also led to irreversible protein unfolding, the next protein sample was heated only to 50°C. The next experiment covered the same temperature range, but at 50°C the sample was incubated for additional 5 min. In the last experiment the protein sample was heated to 55°C, incubated for 5 min and finally cooled to 30°C.

As presented in Fig. 5, the set of data obtained in these five independent experiments covers most of the ellipticity *vs.* temperature diagram.

It must be pointed out that to reduce the influence of sample preparation procedure all the experiments were done using the same stock of the protein. The middle-point temperature of the unfolding process, T_U, was estimated as 86°C, significantly exceeding the apparent value of 55°C estimated from the single 30–80°C unfolding experiment. As appears clearly in the right panel of Fig. 5, the unfolded protein is immediately converted into an irreversible form, F, and as a consequence, even the partial loss of ellipticity cannot be recovered upon cooling (Fig. 5, left panel). This causes the shift of the apparent middle-point temperature as large as approx. 25°C.

The construct of domain 1+2 unfolds at T_{app} = 55.5°C. As for the parent protein, a series of heating–cooling experiments were also performed. Three top temperatures (T_R = 53, 55 and 56°C) were probed, and no incubation at the highest temperature was required. In contrast to the full-length HCV helicase, a partial recovery of the ellipticity was observed upon cooling, with an experimentally estimated rate of 0.5 min⁻¹. This indicates that the rate of transition from the unfolded state, U, to the irreversibly unfolded state, F, is strongly decreased. The average rate, k_f, is approximately three orders of magnitude lower than for the full-length helicase: 7.2×10^{-3} s⁻¹ *vs.* 11 s⁻¹. The estimated parameters of the Lumry–Eyring model indicate that the two-domain construct is thermodynamically less stable than the full-length helicase, with an unfolding middle-point at 56°C and the free energy change of 44 kcal/mol, but in fact the decreased rate of the irreversible unfolding process causes that the construct displays an improved resistance to short-time heating (Fig. 5). Since the refolding rate, U → N, is comparable to the U → F transition rate, partial refolding of the protein could be achieved, whose efficiency depends on the highest temperature reached. Generally, the lower the maximal temperature to which the protein is heated, the larger is the amount of refolded protein.

The separate domain 1 of the HCV helicase is significantly less stable than the full-length helicase

Table 1. Thermodynamic parameters describing thermal unfolding of domain 1 of the HCV helicase placed in different sequential contexts estimated according to the Lumry–Eyring multi-state transition model

	HCV Helicase	Dom. 1+2	Dom. 1	Dom. 2
T_{app} [°C]	54.8 ± 0.3	55.5 ± 0.3	50.8 ± 0.2	82.9 ± 2
T_U [°C]	85.6 ± 2	56.1 ± 1	52.7 ± 1	89.5 ± 5
ΔG^u [kJ/mol]	18.4 ± 3	43.9 ± 3	24.5 ± 1	8.4 ± 0.5
k_R [10^{-3} s⁻¹]	0	7.1	5	0
k_f [10^{-3} s⁻¹]	11500	7.2	7.7	4.1

T_{app} and T_U are the apparent and estimated middle-point temperature of the unfolding, and ΔG^u is the free energy change upon unfolding estimated at middle-point temperature, T_U. The values k_f and k_R are the estimated rates of conversion of unfolded state to the final (F) and folded (N) states, respectively (cf. Eqn. 1). Domain 2 unfolds at the apparent temperature of 83°C when excised from the full-length protein, and remains folded even when domain 1+2 and the entire HCV helicase start to aggregate.
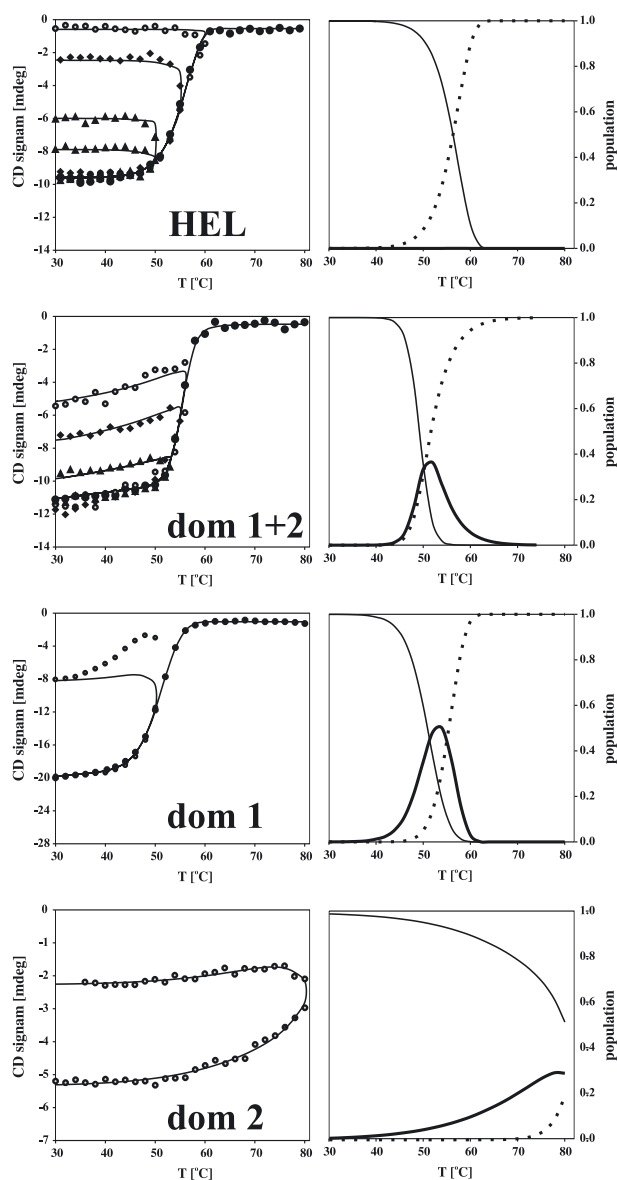
Figure 5. Thermal denaturation of the HCV helicase (HEL), domain 1+2, domain 1, and domain 2 monitored at λ = 220 nm.
Symbols follow experimental data and solid lines represent three-state Lumry-Eyring models fitted (left panels). The estimated partitions of the native (thin line), unfolded (thick line), and irreversibly unfolded (dotted line) states during the basic experiment of heating the samples between 30 and 80°C are presented in the right panels. The parameters of the models are summarized in Table 1.

and domain 1+2, with a transition appearing at T_{app} = 50.8°C. After heating to 50°C and 5 min incubation, the protein refolds with an apparent rate of 0.3 min⁻¹. Unfortunately, no acceptable parameterization of the model was found. The best one, presented in Fig. 5, was used for the rough estimation of the U → F transition rate. This was achieved by neglecting the first five minutes of protein refolding,

and led to a rate of the irreversible step estimated as 8×10^{-3} s⁻¹. The unfolding middle-point was found to be 53°C, close to that of the domain 1+2 construct.

In contrast to domain 1, domain 2 is significantly more stable, even more than the full-length HCV helicase, and up to 50°C (Fig. 5) no thermal unfolding is observed. A characteristic shape of the changes in ellipticity starting at about 60°C is also clearly detectable in the unfolding transitions of the full-length helicase and domain 1+2. This unambiguously demonstrates the multidomain character of the HCV helicase. The dominating transition appears at T_{app} = 83°C. The cooling procedure showed no recovery of ellipticity. Since an additional experiment, performed with a 2-fold decreased concentration of domain 2, displayed no significant changes (not shown), it must be concluded that the final, irreversible unfolding is concentration-independent, and thus is not directly coupled with protein aggregation. The analysis performed on the basis of the Lumry–Eyring model estimated the transition temperature, T_U, as approx. 90°C.

Chemical unfolding of the HCV helicase and its domains

In parallel, alternative titration experiments were performed at a constant temperature with the use of GuHCl as a denaturing agent to avoid protein aggregation upon unfolding. Independent experiments performed with partially unfolded protein proved that HCV helicase samples in 2–3 M GuHCl spontaneously aggregate at the rate of 0.03 min⁻¹. Thus all the experiments concerning the full-length HCV helicase were performed within 15 min after the sample preparation, and immediately after GuHCl addition. The construct comprising domains 1 and 2 was more stable, and when partially unfolded, aggregated at the rate of 1.2×10^{-3} min⁻¹. Separate domains 1 and 2 were stable, even when partially unfolded.

Taking into account the aggregation-related restrictions, for each protein five independent experiments consisting of a series of successive injections of 6 M GuHCl were carried out at 0, 1, 2, 3, 4, and 5 M GuHCl. In all cases the changes of molar ellipticity extracted from five separate experiments fell on a continuous curve after correction for protein dilution, justifying the method used (Fig. 6). As an additional control, spectra of each protein were also recorded in 6 M GuHCl.

The experiments carried out with the separate domains led to an estimate of free energy change, ΔG^U, upon unfolding of 28 and 26 kJ/mol for domain 1 and domain 2, respectively (Table 2). Despite the comparable values of free energy stabilization, the two unfolding profiles differ significantly. This
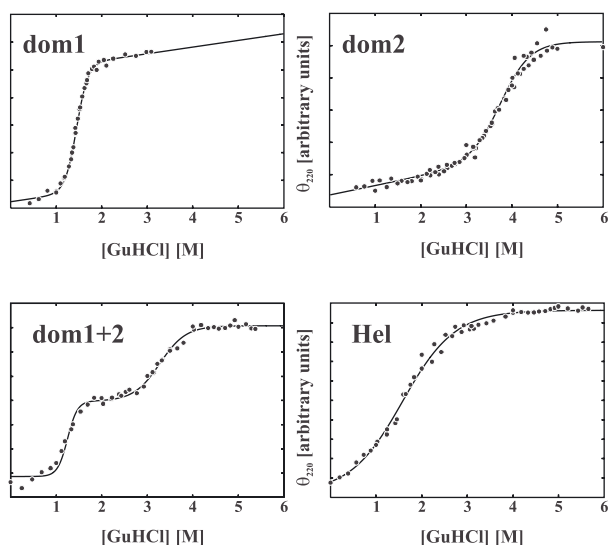
Figure 6. GuHCl-induced denaturation of domain 1, domain 2, domain 1+2 and the HCV helicase (HEL), monitored at λ = 220 nm.

Circles correspond to experimental data, solid lines represent theoretical models. The two-state unfolding processes were modelled with the only exception of domain 1+2, for which the sequence of two two-state transitions was applied. The parameters of the models presented are assembled in Table 2.

results from a difference in the sensitivity to the denaturant, m, 19 and 7 kJ mol^{-1} M^{-1}, respectively. This finding is in agreement with the thermal unfolding analyses, where the denaturation profiles significantly differed with respect to the transition temperature, T_U.

The GuHCl unfolding profile recorded for the construct of domain 1+2 demonstrates the existence of two well-separated transitions, putatively corresponding to sequential unfolding of the two protein domains. More precisely, the transition corresponding to the unfolding of domain 1 is cen-

tered almost entirely in the same region as found for the separate domains (about 1.3 M GuHCl), while domain 2 unfolds slightly more easily in the 1+2 construct, at 3.1 M GuHCl, than when separate, at 3.8 M GuHCl. However, the transition model parameterized, based on the free energy changes estimated for the separate domains, makes it possible to reconstruct the experimental data in an acceptable manner. It is worth pointing out that at high GuHCl concentrations, the mere existence of the unfolded domain 1 causes a slight destabilization of its folded neighbor.

In the case of the full-length HCV helicase, a continuous process of protein denaturation occurs in almost the entire denaturing range used, suggesting that a type of multi-step cooperative unfolding of the helicase domains, additionally coupled with protein aggregation, takes place. The application of models of either sequential or independent transitions generally failed. The only reliable model obtained corresponds to a simple two-state transition, for which the apparent parameters, ΔG^U = 4.4 kJ mol^{-1}, m = 3.3 kJ mol^{-1} M^{-1}, could not be related to any of the previously presented data. Since the full-length helicase tended to insensitively aggregate when partially unfolded, no further analysis was performed.

A comparison of the thermodynamic parameters determined for the four proteins studied provides an insight into the pattern of intramolecular interactions influencing the stability of helicase domains. Thus, domain 2 excised from the context of the full-length protein is stable up to 60–62°C and at 25°C is stable in 3 M GuHCl. The full-length helicase starts to aggregate at 60°C, when its domain 2 is still almost completely folded. This results in an almost linear profile of ellipticity changes in the high temperature region. Domain 1, unfolding separately at T_U = 51°C, is stabilized by interactions with the neighboring domain 2 in both the domain 1+2 con-

Table 2. Thermodynamic parameters describing GuHCl-induced unfolding of the HCV helicase and its domains placed in different contexts

	ΔG^1 [kJ mol^{-1}]	m_1 [kJ mol^{-1} M^{-1}]	C_1 [M]	ΔG^2 [kJ mol^{-1}]	m_2 [kJ mol^{-1} M^{-1}]	C_2 [M]
Helicase	−4.4 ± 1.3	3.3 ± 0.6	1.36			
Dom. 1+2	−14.2 ± 2.2	11.8 ± 1.8	1.21	−46 ± 13	15 ± 4	3.10
	−27.3*	22.6 ± 0.5	1.21	−25.9*	8.7 ± 0.2	2.98
Dom. 1	−27.3 ± 1.7	18.6 ± 1.1	1.47			
Dom. 2				−25.9 ± 2.5	6.9 ± 0.5	3.77

ΔG^i, free energy change; m_i, sensitivity to the denaturant, C_i, GuHCl concentration at the middle-point of unfolding transition; *constrained at the value estimated for the separate domain. The unfolding of domain 1+2 agrees with the superposition of the unfolding phenomena of the separate domains, while the entire helicase exhibits a qualitatively different unfolding profile (Fig. 6). All the values were determined at 25°C.

struct and in the full-length helicase, leading to an increase of its thermal stability by approx. 4°C.

Domain 3 seems crucial for the protein stability. It exhibits the lowest resistance to both the temperature and the denaturant. Its unfolding causes intensive protein aggregation rendering the full-length helicase apparently unstable. Unfortunately, our efforts to experimentally confirm these results by expressing domain 3 in bacteria led to the conclusion that it is toxic for prokaryotic cells.

The domain 1+2 construct is significantly more stable than the entire helicase, exhibiting a tendency towards aggregation decreased by three orders of magnitude. Both the thermal and denaturant stability of this construct are in agreement with the properties of separate domains 1 and 2, and, according to the thermodynamic data (Tables 1 and 2), in both cases the unfolding of domain 1 precedes the denaturation of domain 2. In contrast to the domain 1+2 construct, the entire HCV helicase displays a co-operative-like unfolding profile upon both thermal and chemical denaturation. This clearly indicates that multidomain proteins are complicated systems whose thermodynamic properties are not determined by simple superposition of those estimated for separate domains.

Finally, it must be pointed out that based on their mass, solubility and stability, both domains as well as the two-domain construct adopt secondary structures close to that observed in the full-length protein, and represent good targets for both NMR-based and ITC-monitored drug screening. Our preliminary NMR results, as well as those of Liu and Wyss (2000; 2001) fully support these conclusions.

The CD analysis method presented, by estimating the stability and proper folding of subcloned domains, allows evaluation of the feasibility of further structural studies such as NMR or crystallization even for only partially characterized proteins.

It is worth noting that, in contrast to mechanical stretching (Cieplak & Sulkowska, 2005), the order of chemical denaturation processes agrees with that of thermal unfolding.

Acknowledgements

REFERENCES

Boguszewska-Chachulska AM, Krawczyk M, Stankiewicz A, Gozdek A, Haenni A.-L, Strokovskaya L (2004) Direct fluorometric measurement of hepatitis C virus helicase activity. *FEBS Lett* **567**: 253–258.

Boguszewska-Chachulska AM, Krawczyk M, Najda A, Kopanska A, Stankiewicz A, Zagorski-Ostoja W, Bretner M (2006) Searching for a new anti-HCV therapy: synthesis and properties of tropolone derivatives. *Biochem Biophys Res Commun* **341**: 641–647.

Bohm G, Muhr R, Jaenicke R (1992) Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Eng* **5**: 191–195.

Bretner M (2005) Existing and future therapeutic options for hepatitis C virus infection. *Acta Biochim Polon* **52**: 57–70.

Cieplak M, Sulkowska JI (2005) Thermal unfolding of proteins. *J Chem Phys* **123**: 194–908.

Frost AA, Pearson RG (1961) Transition-state theory: comparison of collision and transition state theories. In *Kinetics and Mechanism*. Frost AA, Pearson RG, eds, pp 93–99. 2nd ed. John Wiley and Sons, New York, USA.

Gesell JJ, Liu D, Madison VS, Hesson T, Wang YS, Weber PC, Wyss DF (2001) Design, high-level expression, purification and characterization of soluble fragments of the hepatitis C virus NS3 RNA helicase suitable for NMR-based drug discovery methods and mechanistic studies. *Protein Eng* **14**: 573–582.

Gordon CP, Keller PA (2005) Control of hepatitis C: a medicinal chemistry perspective. *J Med Chem* **48**: 1–20.

Hadziyannis SJ Jr, Sette H, Morgan TR, Balan V, Diago M, Marcellin P, Ramadori G, Jr Bodenheimer H, Bernstein D, Rizzetto M, Zeuzem S, Pockros PJ, Lin A, Ackrill AM, PEGASYS International Study Group (2004) Peginterferon-alpha2a and ribavirin combination therapy in chronic hepatitis C: a randomized study of treatment duration and ribavirin dose. *Ann Intern Med* **140**: 346–355.

Kelly SM, Price NC (1997) The application of circular dichroism to studies of protein folding and unfolding. *Biochem Biophys Acta* **1338**: 161–185.

Kim DW, Gwack Y, Han JH, Choe J (1995) C-terminal domain of the hepatitis C virus NS3 protein contains an RNA helicase activity. *Biochem Biophys Res Commun* **215**: 160–166.

Kwong AD, Kim JL, Lin C (2000) Structure and function of hepatitis C virus NS3 helicase. *Curr Top Microbiol Immunol* **242**: 171–196.

Lam AM, Keeney D, Frick DN (2003) Two novel conserved motifs in the hepatitis C virus NS3 protein critical for helicase action. *J Biol Chem* **278**: 44514–44524.

Levin MK, Patel SS (2002) Helicase from hepatitis C virus, energetics of DNA binding. *J Biol Chem* **277**: 29377–29385.

Liu D, Wyss DF (2000) Sequence-specific ^{1}H, ^{15}N and ^{13}C resonance assignments for an engineered arginine-rich domain of the hepatitis C virus NS3 RNA helicase. *J Biomol NMR* **18**: 279–280.

Liu D, Wyss DF (2001) Backbone ^{1}H, ^{15}N and ^{13}C resonance assignments of the NTPase subdomain of the hepatitis C virus NS3 RNA helicase. *J Biomol NMR* **19**: 283–284.

Liu D, Windsor WT, Wyss DF (2003) Double-stranded DNA-induced localized unfolding of HCV NS3 helicase subdomain 2. *Protein Sci* **12**: 2757–2767.

Lumry R, Eyring H (1954) Conformation changes of proteins. *J Phys Chem* **58**: 110–120.

Marquardt DW (1963) An algorithm for least-squares estimation of non-linear parameters. *J Soc Ind App Math* **11**: 431–441.

Rosenberg S (2001) Recent advances in the molecular biology of hepatitis C virus. *J Mol Biol* **313**: 451–464.

Sarver RW, Rogers JM, Stockman BJ, Epps DE, DeZwaan J, Harris MS, Baldwin ET (2002) Physical methods to determine the binding mode of putative ligands for hepatitis C virus NS3 helicase. *Anal Biochem* **309**: 186–195.

Williams T, Kelley C (2007) *Gnuplot version 4.1*, Copyright (C) 1986–2007.

Yao N, Hesson T, Cable M, Hong Z, Kwong AD, Le HV, Weber PC (1997) Structure of the hepatitis C virus RNA helicase domain. *Nat Struct Mol Biol* **6**: 463–467.

Zagulski M, Kressler D, Becam AM, Rytka J, Herbert CJ (2003) Mak5p, which is required for the maintenance of the M1 dsRNA virus, is encoded by the yeast ORF YBR142w and is involved in the biogenesis of the 60S subunit of the ribosome. *Mol Genet Genomics* **270**: 216–224.